# Comparing Self-Reported and Partnership-Inferred Sexual Orientation in Household Surveys

Kühne, Simon; Kroh, Martin; Richter, David

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

**Empfohlene Zitierung / Suggested Citation:**
Kühne, S., Kroh, M., & Richter, D. (2019). Comparing Self-Reported and Partnership-Inferred Sexual Orientation in Household Surveys. *Journal of Official Statistics*, *35*(4), 777-805. https://doi.org/10.2478/jos-2019-0033

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 777–805, http://dx.doi.org/10.2478/JOS-2019-0033

# Comparing Self-Reported and Partnership-Inferred Sexual Orientation in Household Surveys

*Simon Kühne*[1], *Martin Kroh*[1], *and David Richter*[2]

Research comparing heterosexuals with bisexuals and homosexuals in economics and the social sciences typically relies on two strategies to identify sexual orientation in existing survey data of general populations. Probing respondents to self-report their sexual orientation is generally considered the preferred option. Since self-reports are unavailable in most large multidisciplinary surveys, often researchers infer sexual orientation from the gender-constellation of a respondent's partnership instead. Based on German Socio-Economic Panel (SOEP) data, this article reviews both strategies empirically in the context of a household panel survey. The analysis shows that self-reported and partnership-inferred sexual orientation are not mutual substitutes, instead leading to substantively different conclusions about differences between heterosexuals and LGBs (Lesbian, Gays, and Bisexuals). The article discusses problems of non-coverage in partnership-inferred sexual orientation and also investigates measurement error in self-reported sexual orientation, finding notable mode and interviewer effects.

*Key words:* Surveys; sexual orientation; measurement error; interviewer effects; survey methodology.

## 1. Motivation

The concept of sexual orientation received not only increasing public, but also academic attention in the past decades. While early scholarly proponents of research on sexual orientation often came from clinical psychology, public health research, and social psychology, the concept has been increasingly adopted by quantitatively oriented scholars from economics and social scientists alike. Sexual orientation is thereby acknowledged to be an important dimension of inequality, structuring societies at large and affecting individuals' lives comprehensively similar to the inequality dimensions of race, gender identity, age, and class. The emergence of the academic field of *Queer Studies* reflects this view that the concept of sexual orientation is multidisciplinary in nature. The cross-cutting nature of the concept as well as its relevance for various disciplines constitute the need for a measurement to become an established part of questionnaires in multidisciplinary surveys, similar to the other inequality dimensions.

The concept of sexual orientation is conceived of by many scholars as a durable sexual attraction to either the different sex (heterosexuals), the same sex (homosexuals), both

[1] Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany. Emails: simon.kuehne@uni-bielefeld.de and martin.kroh@uni-bielefeld.de
[2] Socio-Economic Panel Study (SOEP), DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany. Email: drichter@diw.de
**Acknowledgments:** We would like to thank the Guest Editors Stephanie Steinmetz, Mirjam Fischer, and Nancy Bates, as well as the anonymous reviewers for their valuable comments and time and effort devoted to improving the quality of the article.

sexes (bisexuals), or none of them (for an overview of the debates on the conceptualization of sexual orientation, see, for example, Williams Institute 2009; Haseldon and Joloza 2009; Moliero and Pinto 2015). The term sexual orientation is thereby used rather broadly compared to the more specific concepts of sexual attraction ("I feel attracted to. . ."), sexual behavior ("I had sexual experience with . . ."), and sexual identity ("I am heterosexual, homosexual, . . ."). For instance, Geary et al. (2018) describe sexual attraction, behavior, and identity as subdimensions constituting sexual orientation (see also Chandra et al. 2013; Galupo et al. 2016; Wolff et al. 2016). They show that these facets of sexual orientation do not fully overlap empirically, nor do they serve the same research interests: public health research may be interested more strongly in aspects of sexual behavior, clinical psychology may focus more on emotional attraction, and economics and social sciences more on the social identity aspect of sexual orientation.

While topical surveys often include multiple-item measures of different facts of sexual orientation in their questionnaires, the paucity of any survey items on sexual orientation in multidisciplinary surveys is astonishing. An inquiry of codebooks of the Inter-university Consortium for Political and Social Research (ICPSR) data repository suggests that in only 90 out of the 10,443 archived surveys, that is, less than 1%, do study descriptions include the key words "sexual orientation". Although alternative search terms, such as "lesbian" (525 studies), "gay" (1,873 studies), "bisexual" (356 studies), and "homosexual" (912 studies) lead to higher incidences, these studies often include items on the acceptance of sexual minorities in the general population, rather than identifying respondents by sexual orientation. Browsing these lists suggests that the larger the sample size and the broader the scope of surveys thematically, the less likely that they include measures of sexual orientation. Hence, researchers interested in utilizing existing social surveys, election studies, and census data for research on sexual orientation will either evade to smaller topical studies or they will draw on alternative strategies to identify sexual orientation in respondents.

One popular alternative strategy to directly probing sexual orientation is to indirectly infer an indicator of sexual orientation in existing multidisciplinary surveys from the reported gender-constellation of respondent's current and previous partnership(s). For instance, in household surveys (e.g., the Panel Study of Income Dynamics (PSID), the Survey of Health, Ageing and Retirement in Europe (SHARE), the Socio-Economic Panel (SOEP) – utilized in the present article – and Census Studies of the United States, as well as many other countries), researchers can draw on self-reported gender and the self-reported relationship between two respondents of a household in order to identify same-sex and opposite-sex couples. Although the prevalence of empirical studies on sexual orientation using partnership-inferred measures of sexual orientation clearly varies across disciplines and is low in public health research and queer studies in general, it represents the predominant empirical strategy in other fields. Highly valuable research on poverty rates in sexual minorities, occupational segregation, and the pay gap between heterosexuals and homosexuals relies in large parts on partnership-inferred sexual orientation based on Census data and data from the General Social Survey, for instance (see an overview, Klawitter 2014).

As survey-based research comparing heterosexuals with LGBs is often constrained to using gender constellations of partnership as a substitute for direct measures of sexual

orientation, findings rely on the assumption of the functional equivalence of alternative ways to measuring sexual orientation. This article aims at comparing both strategies of measuring sexual orientation in survey-based research: a) self-reported sexual identity, and b) a partnership-inferred indicator of sexual orientation. For our analysis, we rely on Socio-Economic Panel Study data (SOEP), see Goebel et al. (2018). The SOEP is a nationally representative longitudinal survey of about 15,000 private households in Germany with annual interviews since 1984. All members of the selected households aged 18 years and older are asked to participate in annual interviews. The SOEP consists of multiple subsamples in order to maintain a reasonable panel size and representativeness of the German population over time. In addition to general population subsamples, some subgroups are specifically boosted, including migrants (samples B and M) and families (sample L), see Kroh et al. (2018). Moreover, new household members (e.g., new partners or grown-up children) are invited to join the study. To minimize attrition, individuals are followed even if a household splits or moves. On average, a respondent in wave 2016 already participated for 11 years in the SOEP (Min: 1, Max: 33, Median: 7).

The household structure of the SOEP permits studying partnership-inferred sexual orientation from its first year of interviewing in 1984. Moreover, in 2016, self-reported sexual orientation was surveyed for the first time by SOEP. More specifically, SOEP respondents report their sexual identity by categorizing themselves as heterosexuals, homosexuals, or bisexuals.

Before comparing both strategies of measuring sexual orientations in Section 4, we review the advantages and disadvantages of the two approaches separately. Section 2 discusses self-reported sexual orientation with a focus on measurement error issues. More precisely, we investigate mode and interviewer effects on self-reported sexual orientation and item nonresponse. Section 3 examines partnership-inferred sexual orientation and discusses potential measurement errors such as misclassification. Section 4 firstly investigates the extent to which the two strategies generate the same classification of respondents with respect to sexual orientation. Secondly, we examine differences in the sub-populations classified by each method across a variety of socio-demographics and traits. Finally, Section 5 summarizes the results and offers practical implications for researchers conducting or relying on surveys.

## 2. Self-Reported Sexual Orientation

Different items have been developed in order to measure sexual orientation by self-reporting (Sell 2007; Gates 2011; Wolff et al. 2016). In the following paragraphs, we will focus on the social identity dimension of sexual orientation. Reviewing the existing literature, the most common approach to obtain self-reports on the identity facet of sexual orientation in (large scale) surveys is to outright ask respondents whether they identify themselves as heterosexual, homosexual, or bisexual (e.g., Wooden 2014; Uhrig 2014).

In the 2016 wave, for the first time, respondents of the SOEP were asked about their sexual orientation. The question is worded, “In the context of relationships, the question of sexual orientation arises. Would you describe yourself as...?” The available answers were “Heterosexual or straight (that is, attracted to the opposite sex),” “Homosexual (gay or lesbian, that is, attracted to the same sex),” “Bisexual (attracted to both sexes),” “Other,”

*Table 1. Distribution of responses to the question on self-reported sexual orientation.*

| Response | Percent (unweighted) |
|---|---|
| Heterosexual | 86.11 |
| Homosexual | 0.65 |
| Bisexual | 0.54 |
| Other | 7.10 |
| Prefer not to say | 4.41 |
| No answer (nonresponse) | 1.19 |

n = 24,287.

"Prefer not to say" and "No answer (nonresponse)". A similarly worded question is used by the UK Understanding Society Study (Booker et al. 2017) and the Australian Household, Income and Labour Dynamics Survey (HILDA, Wooden 2014). The explanations in parentheses were added, as earlier pretesting showed that the terms homosexual and heterosexual alone caused misunderstandings among some respondents. Table 1 displays the distribution of responses.

About 1.2% of respondents report identifying as either homosexual or bisexual. Applying cross-sectional weights, which compensate for the disproportionate sampling, as well as nonresponse (Kroh et al. 2018), we estimate that the share of adults in Germany who openly identify as lesbian, gay, or bisexual is 1.9%.

The high number of answers avoiding self-categorization on the dimension of sexual orientation reported in Table 1 shows that, in line with previous research, the strategy of directly probing sexual orientation is not without problems.

A number of previous studies analyzing measurement error in questions about sexual orientation, for instance, showed that a particularly high number of respondents chooses to not answer the question at all (*item nonresponse*), either by outright refusing to answer or by selecting the answer "Other" (e.g., Jans et al. 2015). This response behavior may reflect uncrystallised views on the self-identification, with sexual orientation on the one hand, and the perception of information on sexual orientation as sensitive on the other hand. Incorrect information may be provided intentionally by respondents in order to meet presumed societal expectations (*social desirability bias*, see Krumpal (2013) for an overview) and to obtain social approval (from the interviewer). Moreover, some interviewers may try to avoid a seemingly awkward interview situation by pre-quoting the item nonresponse option. Both social desirability bias and interviewer behavior point to the importance of contextual factors of the interview situation (e.g., Tourangeau and Yan 2007; Kühne 2018; Hilgert et al. 2016).

In the upcoming section, thus, we investigate possible effects of the mode of data collection, and the interviewer, on response behavior to the direct question on sexual orientation in the SOEP. Our analysis aims to provide guidance for researchers who plan to collect data on self-reported sexual orientation of respondents in surveys.

### 2.1. *Mode Effects*

An important choice that researchers make in designing a survey is the mode of data collection. As a multi-mode survey, SOEP allows testing of the effects of interviewer-

administered interviewing versus self-administered interviewing in surveying self-reported sexual orientation.

In the SOEP, data collection is largely based on personal, face-to-face interviews. Since 1998, the SOEP has been gradually replacing paper and pencil interviewing (PAPI) with computer-assisted personal interviewing (CAPI) as the predominant mode of data collection. For instance, in the 2016 wave, 72% of the more than 24,000 respondents who were asked for their sexual orientation were interviewed by an interviewer face-to-face, with 90% of those interviews conducted via CAPI. The remaining 28% of respondents in the 2016 wave used a self-administered mode based on a printed or digital version of the questionnaire and without an interviewer present.

Numerous studies have shown that the mode of data collection is decisive for measurement error in survey questions (e.g., Dillman et al. 2009; Schwarz et al. 1991). For instance, item nonresponse rates tend to be lower in interviewer-administered surveys. There are three main reasons for this. First, in cases where respondents have difficulties in understanding a question, interviewers can clarify questions and answer options, thus helping respondents to provide a valid answer. Second, interviewers may actively probe and argue in order to obtain valid responses (Kuha et al. 2014). Third, many respondents likely perceive that not providing an answer is an undesired behavior in front of the interviewer because it runs counter to the main purpose of the interview of collecting valid information.

These advantages of face-to-face interviewing suggest that respondents are more likely to provide responses in interviewer-administered interviewing on the one hand. However, on the other hand, the privacy of the self-interview may have positive effects on measurement, as respondents may feel more comfortable divulging sensitive information on sexual orientation.

Table 2 displays responses to the sexual orientation question across modes. 4.41% of respondents refused to answer the question by actively stating “Prefer not to say”. With 5.40%, the share is higher in the Mail/CAWI (Computer-Assisted Web Interview) mode compared to interviewer-administered interviews (4.02%). Moreover, in the Mail/CAWI mode, 3.29% did not provide any answer at all (nonresponse). Missing information in the CAPI mode amounts to 0.38% only – most likely due to the fact that the interviewers manage the CAPI system and aim at low nonresponse rates. In total, 5.60% of all respondents choose to not provide a valid answer, with 8.69% in Mail/CAWI and 4.40% in the CAPI mode. Item nonresponse rates in nearby questions in the questionnaire are much lower, suggesting that the observed response behavior is specific to the question on sexual orientation and does not reflect a general tendency of respondents to not provide valid answers (e.g., as a type of satisficing behavior, see Krosnick et al. 1996). The observed item nonresponse rate is quite high compared to other existing studies surveying sexual orientation (3.4% [Wooden 2014], 3.2% [Uhrig 2014], 2.78% [Frederiksen-Goldsen and Kim 2014], 0.93% [Dahlhamer et al. 2014]).

The presence of an interviewer may not only affect the propensity of item nonresponse, but also the selection of valid answer options. In fact, socially desirable response behavior and impression management in interviewees tend to be more prevalent when interviewers (and others) are present (Krumpal 2013). Social desirability bias is shown to be most prevalent in questions that are perceived to be sensitive and stigmatizing (e.g., Tourangeau

*Table 2. Distribution of self-reported sexual orientation across survey modes.*

| Response | F2F % | Mail/CAWI % | Total % |
|---|---|---|---|
| Heterosexual | 87.46 | 82.60 | 86.11 |
| Homosexual | 0.56 | 0.88 | 0.65 |
| Bisexual | 0.36 | 1.02 | 0.54 |
| Other | 7.21 | 6.81 | 7.10 |
| Prefer not to say | 4.02 | 5.40 | 4.41 |
| No answer (nonresponse) | 0.38 | 3.29 | 1.19 |

In percent. Unweighted. n = 24,287.

and Yan 2007), such as sexual orientation for lesbian, gay, or bisexual respondents. Some may experience difficulties in disclosing their sexual orientation to another person, particularly in cases where (parts of) the social environment, such as other household members, neighbors, and colleagues are unaware of the respondent's sexual orientation. Hence, we expected a lower share of respondents reporting an LGB sexual orientation in the interviewer-administered mode than in the self-administered mode, which is perceived as more anonymous. The results in Table 2 match our expectations, with 1.90% reporting a homosexual or bisexual orientation in the self-administered, compared to 0.92% in the interviewer-administered survey mode.

Related to this, an unexpected result is the large share of respondents choosing the answer category "Other" (7.10%). We initially implemented this category for respondents that do not identify as either heterosexual, homosexual, or bisexual, but prefer other forms of sexual self-categorization, such as asexual, pansexual, and queer. While there is no reliable population estimate of the number of, for example, persons identifying as asexual and pansexual available for Germany, Aicken et al. (2013) estimate the prevalence of asexuals in the UK at 0.4%. This is supported by results of the SOEP Innovation Sample in 2015 (see Richter and Schupp 2015) in which only 0.16% of over 5,000 respondents reported an asexual identity.

Hence, it is unlikely that the 7.10% of respondents who choose the "Other" category represent asexuals, for instance. Although follow-up write-ins for those who report "Other" have not been collected in the main survey, we surveyed this information in 2006 in a pretest of 1,057 respondents. Here, only eight respondents provided write-ins and none of them reported a queer sexual orientation, such as asexual or pansexual; rather, on the contrary, seven mentioned variations of "normal" and a single respondent wrote "I don't know these words". Hence, we interpret the high number of respondents reporting the "Other" category in most cases as an unanticipated form of item nonresponse to the sexual orientation question (including 'refusal', 'don't know', and 'does not apply').

In Table 3, we report the results of multilevel logistic regression models with respondents nested in households and households nested in interviewers assessing mode effects on response behavior. Although interviewers do not actively collect information in the self-administered mode, they may be involved, nonetheless, in the process of contacting these households and thus promoting participation. Thus, we use the multilevel data structure with respondents nested in interviewers even for households with self-administered interviews. Hierarchical regression modelling allows to address the

*Table 3. Mode effects on response propensity, reports of homosexual and bisexual orientation, and reporting "other".*

| | Response (yes/no) | | Homosexual/bisexual \| response = yes | | Reporting "other" \| response = yes | |
|---|---|---|---|---|---|---|
| Variable | OR | 95% CI | OR | 95% CI | OR | 95% CI |
| Mode | | | | | | |
| Face to face (Ref.) | – | – | – | – | – | – |
| Mail/CAWI | 0.38*** | 0.29– 0.49 | 4.60*** | 2.37–8.89 | 2.79*** | 1.64–4.72 |
| Third person present | | | | | | |
| No third person (ref.) | – | – | – | – | – | – |
| Partner | 0.76* | 0.59–0.97 | 0.77 | 0.38–1.56 | 1.27 | 0.82–1.96 |
| Other person | 0.71* | 0.51–0.98 | 2.83** | 1.34–5.96 | 1.51 | 0.82–2.78 |
| No information | 0.41*** | 0.30–0.56 | 2.92* | 1.25–6.84 | 3.69*** | 2.10–6.48 |
| Respondent panel experience | | | | | | |
| 1–2 years (ref.) | – | – | – | – | – | – |
| 3–5 years | 1.84** | 1.24–2.73 | 0.87 | 0.40–1.88 | 0.71 | 0.33–1.57 |
| 6–10 years | 2.35*** | 1.52–3.63 | 0.83 | 0.37–1.83 | 0.90 | 0.39–2.05 |
| 11+ years | 1.12 | 0.72–1.75 | 0.66 | 0.27–1.64 | 2.40 | 0.99–5.78 |
| Subsample | | | | | | |
| Cross–sectional (ref.) | – | – | – | – | – | – |
| Migrants | 0.70 | 0.48–1.00 | 0.18** | 0.07–0.48 | 1.03 | 0.49–2.18 |
| Families | 1.41 | 0.96–1.07 | 0.81 | 0.39–1.66 | 2.29* | 1.14–4.57 |
| Other | 0.92 | 0.66–1.28 | 0.32* | 0.12–0.83 | 0.62 | 0.32–1.20 |
| Respondent gender | | | | | | |
| Male (ref.) | – | – | – | – | – | – |
| Female | 0.80** | 0.68–0.93 | 1.30 | 0.87–1.94 | 1.54*** | 1.20–1.97 |
| Respondent age | | | | | | |
| <29 (ref.) | – | – | – | – | – | – |
| 30–39 | 1.33 | 0.91–1.94 | 0.38** | 0.18–0.78 | 0.96 | 0.48–1.90 |
| 40–49 | 0.88 | 0.63–1.25 | 0.28*** | 0.14–0.57 | 1.44 | 0.77–2.69 |

*Table 3. Continued.*

| Variable | Response (yes/no) | | Homosexual/bisexual \| response = yes | | Reporting "other" \| response = yes | |
|---|---|---|---|---|---|---|
| | OR | 95% CI | OR | 95% CI | OR | 95% CI |
| 50–59 | 0.55** | 0.40–0.78 | 0.12*** | 0.06–0.30 | 3.09*** | 1.64–5.79 |
| 60–69 | 0.43*** | 0.29–0.62 | 0.06*** | 0.02–0.17 | 14.41*** | 6.88–30.17 |
| 70–79 | 0.28*** | 0.19–0.42 | 0.03*** | 0.01–0.12 | 51.17*** | 22.28–117.57 |
| 80+ | 0.18*** | 0.11–0.29 | 0.01*** | 0.00–0.04 | 67.48*** | 25.15–181.08 |
| In a relationship? | | | | | | |
| No (ref.) | – | – | – | – | – | – |
| Yes | 2.89*** | 2.26–3.65 | 0.87 | 0.52–1.46 | 0.28*** | 0.18–0.43 |
| Household size | | | | | | |
| Single (ref.) | – | – | – | – | – | – |
| 2 | 1.18 | 0.82–1.66 | 0.53 | 0.25–1.13 | 1.27 | 0.61–2.64 |
| 3 | 1.12 | 0.76–1.64 | 0.24*** | 0.06–0.35 | 4.24*** | 1.93–9.32 |
| 4 | 0.88 | 0.61–1.30 | 0.05*** | 0.02–0.14 | 3.97** | 1.78–8.86 |
| 5 | 0.74 | 0.48–1.13 | 0.05*** | 0.02–0.16 | 2.60* | 1.07–6.32 |
| 6+ | 0.70 | 0.44–1.10 | 0.05*** | 0.02–0.20 | 1.80 | 0.70–4.67 |
| German region | | | | | | |
| North (ref.) | – | – | – | – | – | – |
| East | 0.57* | 0.32–0.99 | 1.09 | 0.44–2.69 | 0.51 | 0.14–1.82 |
| South | 0.55* | 0.32–0.96 | 0.36* | 0.14–0.93 | 0.68 | 0.19–2.40 |
| West | 062 | 0.36–1.05 | 0.77 | 0.33–1.82 | 0.63 | 0.19–2.14 |
| Constant | 0.00*** | 0.00–0.01 | 0.00*** | 0.00–0.01 | 0.00*** | 0.00–0.00 |
| $n_{resp}$ | 24,287 | | 22,928 | | 22,928 | |

*p < .05, **p < .01, ***p < .001.

Multilevel logistic regression with respondents nested in households and households nested in interviewers.

See Table A1 in the Appendix (Section 6) for a comparison of characteristics across modes.

hierarchical nature of the survey data at the level of respondents due to specifics of the sampling and survey design, and thereby improves the estimation of coefficients and their standard errors. We estimate mode effects on three outcomes independently. First, whether there is a valid response at all (yes/no) and, second, whether a homosexual or bisexual orientation is reported (yes/no), conditional on response, and third, whether "Other" is reported (yes/no), conditional on response. Please note that in the SOEP, the assigned mode is not randomly allocated across households and respondents. Thus, we add a number of potential confounding factors to each model. The most influencing factor of allocating mode is the predominant field work routine in a subsample of SOEP. For instance, from wave 2011 onwards, all refreshment and enlargement samples are exclusively interviewed in the CAPI mode. We control for predominant field work routine by adding a subsample identifier into the analysis. Other factors that correlate with mode in the SOEP are respondent's gender, age, and years in the panel. Moreover, we control for correlating household characteristics including whether another person was present in an interview, household size and geographic region in Germany. Finally, we control for current relationship status (yes/no). As an additional robustness check, we replicated our analyses for each subsample and separated for paper-and-pencil and computer-assisted mode. No substantial differences are observed.

The first model estimates the effect of mode on response propensity (response versus item nonresponse). Compared to the face-to-face mode (reference), and in line with the results in Table 2, the odds of providing a valid response in the self-administered mode are less than half of the odds in the interviewer-administered mode (Odds Ratio (OR) = 0.38, $p < .001$). In addition, there are a number of other interesting results when turning to the respondent level characteristics. We find, for instance, lower response propensities for females and older respondents. Moreover, those currently in a relationship are more likely to respond. A possible explanation could be that some respondents understand the survey question as only relating to sexual orientation in a currently ongoing relationship, rather than general sexual orientation. This seems plausible, as the question was placed right after questions about family and relationship status.

In the second model, we estimate mode effects on the probability of reporting a homosexual or bisexual orientation given a valid response. Respondents participating via Mail/CAWI are associated with a 4.60 times higher chance (odds) of reporting a homosexual or bisexual orientation ($p < .001$) rather than a heterosexual orientation or answering "Other". Again, the results match expectations for reporting a homosexual or bisexual orientation; respondents are much more likely to report sensitive and potentially undesired responses in the self-administered – and likely more private – interview mode. In addition, effects of respondent-level characteristics match expectations. The propensity of reporting a homosexual or bisexual orientation decreases dramatically with the increasing age of respondents. This is plausible, as older cohorts are less likely to openly identify and live as non-heterosexual. Moreover, the larger the household size, the less likely it is to obtain a non-heterosexual response. This is most likely due to the fact that LGBs more often live in smaller households compared to heterosexual households. Note that our estimation is associated with comparatively large uncertainty, as only few respondents report a non-heterosexual relationship; as a consequence, the 95% confidence interval for the estimated odds ratio is comparatively wide.

In the third model, we estimate mode effects on respondents selecting the "Other" category as an answer. The odds ratio of respondents using the "Other" option in self-administered versus interviewer-administered mode is estimated at 2.79 ($p < .001$). Hence, in persons who do not explicitly refuse an answer on sexual orientation, it is much more likely that they choose "Other" when not being interviewed by an interviewer in person. Turning to respondent level characteristics, and matching the results of Model 1, female and older respondents are more likely to respond "Other".

The latter two models reported in Table 3 on homosexual and bisexual responses as well as the "Other" response option consider the first stage of the answering process, that is, refusal or response, to be uninformative for the second stage. It may very well be, however, that both stages of the response process are correlated. Heckman selection models allow modelling correlated residuals at stages one and two (Heckman 1976, 1979, Puhani 2000). Estimating Heckman models (probit, not reported in the form of a table) suggest that the first stage of response versus refusal and the second stage of LGB response are only moderately correlated and thus, estimates of Table 3 are hardly affected by a change in model specification. However, the first stage of response versus refusal and the second stage "Other" option are positively and strongly correlated, suggesting that some respondents use the "Other" option as a substitute for directing refusing a response. Also, the Heckman specification suggests that the effect of self-administered interviewing turns negative for the "Other" option; most other estimated remain largely unchanged. The change of the direction of the mode effect suggests that interviewer-administered interviews increase refusals both in the form of direct refusals and in the form of the hidden refusal using the "Other" option.

To sum up, observing a valid response is more likely in face-to-face interviewing than in self-administered modes. However, more respondents are willing to share a non-heterosexual orientation in the self-administered mode compared to interviewer-administered interviews. The latter inflates the propensity of a hidden item nonresponse using the "Other" response option. This has implications for measuring sexual orientation in surveys. Our results suggest a strategy of relying on face-to-face interviewing in order to minimise item nonresponse, but to switch to computer-assisted self-interviewing approach (CASI) as a more private mode of interviewer-administered data collection.

### 2.2. *Interviewer Effects*

Some interviewers may generally be more efficient in obtaining valid responses than others. Thus, in addition to mode effects, we also investigate interviewer effects (see West and Blom 2017 for an overview) on the prevalence of item nonresponse and the selection of the "Other" category. As there are only 290 respondents that report a non-heterosexual orientation, and these cases are often clustered within households and interviewers, we did not investigate potential interviewer effects (intra-interviewer clustering) in LGB responses.

Consequently, in this section, we only analyze interviewer-administered interviews, which reduces the number of observations by about one third. On average, each SOEP interviewer conducts 68 interviews in wave 2016 (Median $= 59$, Min $= 1$, Max $= 207$). Item nonresponse rates in the question on sexual orientation vary greatly across interviewers. For the upcoming analysis, we only analyze interviewers that conducted at least five personal interviews (416 interviewers). On average, an interviewer obtains 6%

non-valid responses (Median = 0.00, SD = 13.74). While 59% of the interviewers do not collect any invalid information for the question about sexual orientation at all, many of the interviewer staff did, showing large variation with up to 100% item nonresponse rates (two interviewers; $p_{75} = 4.30\%$; $p_{90} = 16.67\%$). Turning to participants reporting "Other" as an answer, the interviewer-average amounts to 9% (Median = 0.00, SD = 21.74). Again, the majority of interviewers did not collect any "Other" answer (64%). However, there are 13 interviewers that exclusively (100%) obtained "Other" as responses ($p_{75} = 4.35\%$; $p_{90} = 26.67\%$). These 13 interviewers are slightly older and more experienced compared to the rest of the interviewer staff. They do not differ in terms of gender and education.

Are these differences across interviewers driven by characteristics of the interviewers or by respondent, household, or other confounding factors? We use the widely accepted intra-interviewer variance coefficient $\rho_{int}$ proposed by Kish (1962) to quantify interviewer variance in item nonresponse. Interviewer variance relates the interviewers' contribution to the total variation in a survey variable, resulting from the individual biases introduced by each interviewer. The more homogeneous the responses collected by individual interviewers, compared across interviewers, the higher the share of variance that is due to the interviewers. There is a large body of literature on measures of intra-interviewer correlation in survey variables (e.g., Groves 2004, 365; Schnell and Kreuter 2005; West and Olson 2010). They show that interviewer effects are present across all survey topics and question types.

Multilevel cross-classified linear mixed models (see Rasbash and Goldstein 1994) are used to estimate the intra-interviewer variance for item nonresponse and the replying of "Other". These models acknowledge that respondents are nested in households and households are nested in a cross-classified structure of geographic areas (German counties) and interviewers. Using cross-classified models can allow separating interviewer from area effects, a general problem in many large-scale surveys as interviewers are allocated to a specific geographic area, that is, single or few sample clusters only. In these cases, estimated interviewer effects are likely confounded with area effects (Schnell and Kreuter, 2005; Campanelli and O'Muricheartaigh, 1999; Durrant and D'Arrigo, 2014); answers observed by a single interviewer may be more homogeneous not because of the interviewers' biasing effects on responses, but due to the homogeneity of individuals living in the same geographic area. In the SOEP, on average, each interviewer is assigned to five German counties (Median = 5, Min = 1, Max = 17). In each county, on average, six interviewers are conducting interviews (Median = 5, Min = 1, Max = 36).

Whether a respondent provides a non-valid, that is, missing answer (y/n), functions as the dependent variable in the first model. In the second model, and similar to the above analysis of mode effects, we estimate the probability of respondents replying "Other" (yes/no). Adding a number of geographic area covariates and respondent characteristics in our models further minimizes potential problems due to confounding with area effects. Controls include respondent socio-demographics (gender, age, education), current partnership status (yes/no), as well as area characteristics at neighborhood level (street type, age distribution, share of Turkish migrants, move turnout, socio-economic status), at municipality level (size, voting results of the 2013 German Federal Election, age distribution), as well as at county level (unemployment rate, share of foreigners, share of higher educated employees, share of students). The intra-class correlation for the

interviewers is then derived as the share of the interviewer-level variance compared to the total variance that is decomposed into interviewer-specific variance, the area-specific variance, the household-specific variance, and the individual respondent residual variance.

The estimated share of variance in item nonresponse (yes/no) that is due to the interviewer is 0.43 (or 43%). Thus, almost half of the variance in item (non)response originates from the interviewer level. In other words, the interviewer largely influences the propensity of a respondent to provide a valid answer or not. A possible explanation is that some interviewers may feel uncomfortable asking the question because they perceive it as too sensitive, and thus, choose to not ask the question at all. With respect to the selection of "Other" as an answer, the interviewer variance is estimated even higher at 0.88%. Thus, almost 90% of the variance in choosing the other category can be explained by knowing which respondent is allocated to which interviewer. A possible explanation could be that some interviewers choose to not read out the answer option "Other" at all. On the other hand, some interviewers may even suggest to respondents to answer "Other" because they perceive other answers as being too sensitive. Moreover, some interviewers may just not read out the question at all and just answer "Other" for the respondent. While this clearly violates the interview protocol, it is usually not detected by fieldwork management as a) the response is not flagged as item "nonresponse", and b) there is no chance for inconsistencies in responses, since "Other" is compatible with any gender and partnership-constellation. Finally, interviewers themselves may misinterpret the question as relating to a person currently in a relationship only.

Do specific interviewer characteristics explain the observed nonresponse rates? For instance, older interviewers may feel less comfortable asking about sexual orientation, thus achieving lower response rates. To test this, we add a number of interviewer characteristics into the models. This includes socio-demographics (gender, age, education), work experience, personality traits (Big Five, see McCrae and John 1992), as well as political attitudes.

Table 4 displays the results of two multilevel logistic regressions. Although interviewers differ considerably in item nonresponse rates, few of the interviewer characteristics tested exert a statistically significant effect on the observed response behavior. Interviewers with many years of experience (21+ years) obtain more item nonresponse compared to interviewers with up to ten years of work experience (OR = 2.82). Moreover, interviewers with a higher workload are more successful in obtaining valid responses (OR = 0.43 and 0.37). In this regard, the workload itself likely does not have a direct effect on responses. Rather, the workload reflects interviewer skills and experience, as more experienced interviewers are usually given higher workloads. Older interviewers collect more "Other" responses (OR = 1.09) while higher educated interviewers achieve less "Other" responses (OR = 0.32, and = 0.31). Finally, there is evidence that the interviewers' self-reported personality is associated with response: interviewers that describe themselves as comparatively extroverted are more likely to obtain "Other" as an answer (OR = 1.61).

## 3. A Partnership-Inferred Proxy of Sexual Orientation

Presumably, the most frequent form of data that researchers apply to operationalizing same-sex and opposite-sex couples is the 'household grid' (or household-matrix). This

*Table 4. Explaining interviewer effects.*

| | Nonresponse (yes/no) | | Reply "other" | |
|---|---|---|---|---|
| Variable (Interviewer) | Odds ratio | 95%–CI | Odds ratio | 95%–CI |
| Gender | | | | |
| Male (ref.) | – | – | – | – |
| Female | 0.80 | 0.47–1.38 | 0.66 | 0.32–1.33 |
| Age | 1.01 | 0.98–1.04 | 1.09*** | 1.05–1.15 |
| Education | | | | |
| Basic (ref.) | – | – | – | – |
| Secondary | 0.70 | 0.34–1.42 | 0.32* | 0.13–0.77 |
| Tertiary | 0.73 | 0.34–1.55 | 0.31* | 0.12–0.81 |
| Work experience (years) | | | | |
| 1–10 (ref.) | – | – | – | – |
| 11–20 | 1.50 | 0.82–2.76 | 0.75 | 0.35–1.62 |
| 21+ | 2.82** | 1.38–5.74 | 0.57 | 0.21–1.58 |
| Number of personal interviews in wave 2016 | | | | |
| 1–49 | – | – | – | – |
| 50–99 | 0.43** | 0.24–0.77 | 0.69 | 0.33–1.46 |
| 100+ | 0.37* | 0.17–0.79 | 0.94 | 0.36–2.40 |
| Personality score (Big five)[1] | | | | |
| Openness | 1.04 | 0.79–1.35 | 1.12 | 0.78–1.64 |
| Conscientiousness | 1.37 | 0.96–1.95 | 1.41 | 0.89–2.21 |
| Extraversion | 0.91 | 0.65–1.26 | 1.61* | 1.03–2.50 |
| Agreeableness | 0.98 | 0.70–1.39 | 1.42 | 0.91–2.21 |
| Neuroticism | 1.19 | 0.96–1.49 | 1.11 | 0.84–1.48 |
| Party identification | | | | |
| No party (ref.) | – | – | – | – |
| SPD | 0.72 | 0.33–1.57 | 0.62 | 0.22–1.75 |
| CDU/CSU | 0.59 | 0.31–1.14 | 1.47 | 0.65–3.34 |
| Greens | 0.65 | 0.23–1.80 | 0.37 | 0.08–1.67 |
| The left | 0.44 | 0.15–1.35 | 1.71 | 0.44–6.71 |
| Other | 0.76 | 0.24–2.52 | 0.98 | 0.21–4.48 |
| Constant | 0.00** | 0.00–0.24 | 0.00*** | 0.00–0.00 |
| $n_{respondents}$ | 14,521 | | 13,389 | |
| $n_{interviewers}$ | 432 | | 401 | |

*p < .05, **p < .01, ***p < .001.

[1]Range: 1 to 5 with lower values representing less open, agreeable, etc.

Multilevel logistic regression with respondents nested in interviewers.

Controls: Respondent gender, age, education, years in the panel, geographic region (north, east, south, west), whether another person was present (none, partner, other, no information), whether a respondent was in a relationship (y/n), household size, municipality size as well as an initial subsample identifier.

See Table A2 in the Appendix (Section 6) for the full model coefficients.

grid, which is typically used in census surveys, enlists all residents of a household. The responding householder reports, among others, gender of other household members and their relationship to these persons (e.g., partner, mother, child). In some other cases, questionnaires include items on the gender of partners, irrespective of whether the respondent and the partner currently cohabitate. In addition, longitudinal household surveys (e.g., SOEP, PSID, and Understanding Society) provide not only information on current partnerships, but also on past relationships, as long as they fall into the period of observation. In principle, biographical questionnaires implemented in longitudinal, as well as cross-sectional, surveys allow an extension of this period to any prior partnership (Bates and DeMaio 2013). However, a cursory search of partnership biographies in different surveys suggests that only in a minority of cases, the gender of the previous partner was included among the surveyed items. Conversely, in many questionnaires the introductory text presumes that partners are opposite-sex or restricts previous partnership to marriage, which in many countries excludes same-sex partnerships.

### *3.1. Prevalence of Same-Sex Couples*

To provide an overview of the incidence of same-sex couples in the SOEP, Figure 1 compares the proportion of same-sex couples among all couples in the SOEP in 2016 with corresponding estimates from the 2015 German Microcensus (MZ), the 2010 U.S. Census, as well as cross-national estimates of the European Social Survey (ESS), pooling its 2002–2016 waves. The ESS cross-national estimates range from zero same-sex couples in the samples of Poland and Russia to almost 2% of all couples in Denmark. According to the ESS estimates, Germany is among the countries with above-average numbers of same-sex couples in Europe. Figure 1 also displays sizable differences between surveys within Germany. While the 2016 German Microcensus reports only 0.46% same-sex couples and the German ESS samples 1.7%, the SOEP estimate is 0.9% of all couples in Germany.

### *3.2. Using Partnership Information as a Proxy of Sexual Orientation*

Many surveys ask their respondents about partnership and (typically, binary) gender of the partner. This allows distinguishing between respondents not in a partnership, respondents in a partnership with a person of the opposite sex, and respondents in a same-sex partnership.

Researchers use partnership information to infer a proxy of sexual orientation in case a direct measurement is not available (Black 2000; Black et al. 2007, Antecol et al. 2008; Leppel 2009; Klawitter 2011; Liu et al. 2013). This holds true for many household surveys and censuses in particular. In these studies, partnership information is the only information available for analyzing the life of (parts of the) LGB population and comparing them (again, with parts of the) heterosexual population.

There are a number of measurement issues with regard to a partnership-inferred proxy of sexual orientation. Firstly, it does not allow a distinction between a bisexual orientation and a homosexual or heterosexual orientation. A respondent who reports being in a same-sex partnership may be homosexual or bisexual, just like a respondent who reports being in an opposite-sex relationship may be heterosexual or bisexual.

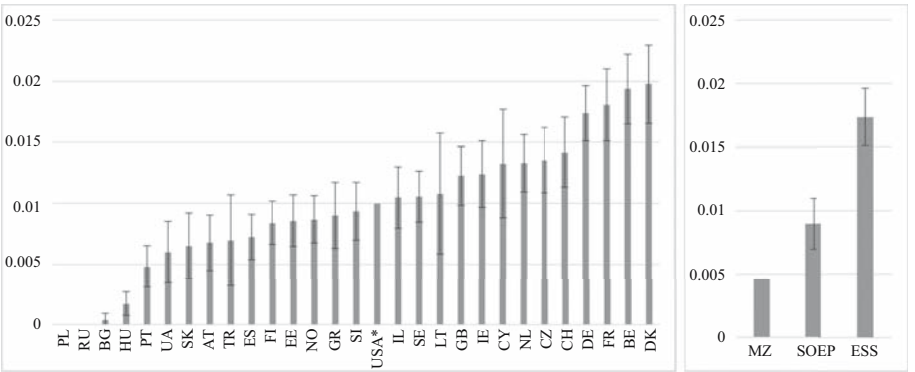


*Fig. 1. Prevalence of same-sex couples across 29 countries (left) and three surveys in Germany (right).*
Note. Population estimates of the rate of same-sex couples among all cohabitating couples.
Source. European Social Survey in waves 2002 to 2016 (Rounds 1–8). Socio-Economic Panel 2016 (SOEP.v33.1), U.S. Census (2010), 2015 German Microcensus (https://www.destatis.de/EN/FactsFigures/InFocus/Population/SameSexCouples.html). For estimates from the German Family Panel (pairfam), see Hank and Wetzel (2018). The estimated prevalence of SOEP uses information on the gender constellation of cohabiting partners only. All estimates weighted applying cross-sectional weights.
*2010 U.S. Census (Lofquist et al. 2012).

Secondly, the strategy relates to couples only. Thus, any comparisons between the LGB and heterosexual sub-populations are restricted to those who state that they are currently in a partnership, which excludes singles and couples living apart together.

Thirdly, respondents may be in a same-sex relationship, but identify as heterosexual and vice versa. As sexual orientation is subjective, the adequacy of 'objective' (observable) characteristics and measures, such as current partnership, is generally limited. Thus, partnership information is a valid proxy for some individuals, but not for others.

Finally, reporting error in the – often complex – survey instrument can result in the misclassification of same-sex and opposite-sex couples. Cortina and Festy (2014) provide an overview of different sources of reporting error, as well as the presence of this problem in census data of different countries (see also Festy 2007). Research on U.S. Census data shows that reporting error considerably inflates the number of same-sex couples. Kreider and Lofquist (2015), for instance, identify misreporting by matching couples from the U.S. Census with Social Security Administration records. A sizable number of U.S. Census same-sex couples appear to be opposite-sex in administrative records. Lengerer (2017), analyzing German Micro Census data, shows that item nonresponse in the household grid is a source of underestimation of the number of same-sex couples. DeMaio et al. (2013), classifying household member names by gender, also find considerable mismatches between reported gender in the household grid and the name-based gender of household members. These studies also identify ways to minimize reporting error by adapting the data collection mode and questionnaire design, for instance (see also Lewis et al. 2015).

## 4. Comparing Self-Reported and Partnership-Inferred Sexual Orientation

As previous research and this article suggest, both self-reported and partnership-inferred sexual orientation are potentially plagued by bias attributable to coverage, as well as

measurement. In the absence of information of the "true" sexual orientation of respondents, we are unable to compare the relative quality of both strategies directly. Thus, the following paragraphs provide some evidence on the extent to which both strategies lead to similar substantive conclusions about the differences between heterosexuals, on the one hand, and homosexuals and bisexuals, on the other hand.

For the partnership-inferred proxy of sexual orientation, we not only use information provided in the SOEP wave 2016, but any partnership-information available in the current and previous waves 2011 to 2015, including partnership-biographies. With this approach, we aim to minimise the risk of misclassifications due to bisexuals incorrectly classified as heterosexuals. In this regard, we refrain from coding sexual orientation if there is only a single piece of information on an opposite-sex partnership. Thus, respondents were classified as LGB in cases where they report a same-sex partnership once in the waves from 2011 to 2016. They are classified as heterosexual if they report an opposite-sex partnership in at least two waves (and never report a same-sex partnership). As a consequence, for many respondents, a classification is not possible because they have not been in a partnership within the observation period (referred to as "no partnership information" in the following).

Table 5 reports a cross-table of the number of SOEP respondents in 2016 classified as LGB by partnership-information, as well as their self-identification. Among the over 24,000 SOEP respondents who participated in 2016, on the basis of partnership-information, we categorize 153 as LGBs and 18,452 as heterosexuals (compared to 290 LGBs and 20,914 heterosexuals based on self-reported sexual orientation). In 5,682 cases (23%), no classification based on partnership-information was possible because a respondent has not been in a partnership over the last years. In comparatively few cases (747; 3%), a respondent provides a non-substantial answer *and* no information on current and previous partnerships.

As expected, for some respondents, the strategies of estimating sexual orientation lead to conflicting results. Thirty-two respondents out of more than 16,000, who self-report as heterosexual, were coded as LGB on the basis of partnership data. We analyzed whether these 32 cases are more recent additions to the panel, thus, pointing to the possibility that respondents may need more time to openly identify as Lesbian, Gay, or Bisexual in a survey. However, they do not differ significantly from the rest of the sample in terms of panel experience. Conversely, 66 out of 290 respondents, that is, 23%, who self-reported being homosexual or bisexual were coded as heterosexuals on the basis of available partnership data. As expected, the majority of these respondents identify as bisexual. Moreover, the conflicting classification results from the fact that both measures capture different aspects of sexual orientation. While the self-reports relate to the self-identification and self-perception of each individual, the partnership-inferred measure reflects acting sexual orientation in relationships – and these aspects can, but do not necessarily match.

Rather surprisingly, only 111 respondents (102 lesbian/gay, 9 bisexual) have consistent information indicating homosexual and bisexual orientation both in partnership and in self-reported data.

What can we learn from these results for survey practice? First of all, the cross-table emphasizes the fact that each method is accompanied by measurement error. While

*Table 5. Self-reported and partnership-inferred sexual orientation.*

| | *Partnership-inferred* | | | |
|---|---|---|---|---|
| *Self-Reported* | Opposite-sex partnership (Heterosexual) | Same-sex partnership (LGB) | No partnership information | Total |
| Heterosexual | 16,060 | 32 | 4,822 | 20,914 |
| Lesbian/gay | 11 | 102 | 45 | 158 |
| Bisexual | 55 | 9 | 68 | 132 |
| Other | 1,337 | 1 | 386 | 1,724 |
| Refused | 989 | 9 | 361 | 1,359 |
| Total | 18,452 | 153 | 5,682 | 24,287 |

Source. SOEP.v33, Wave 2016.

self-reported sexual identity is associated with substantial item nonresponse, partnership-inferred sexual orientation faces problems due to the non-coverage of singles and potential misclassification of bisexual respondents. Second, for many respondents, we find one measure of sexual orientation to substitute missing information on the other measure. Turning to LGBs, in 113 cases, absence of partnership data coincides with homosexual and bisexual self-reports, while in ten cases, absence of self-reported data coincides with same-sex partnership data. Using partnership-inferred sexual orientation may be used to fill in missing information in self-reported sexual orientation. Of the 3,083 respondents who either refused to answer or answered "Other", 2,336 respondents (76%) can be categorized as (likely) heterosexual or LGB, based on their present and previous partnership information. Thus, while some respondents (and their interviewers) may feel uncomfortable talking directly about their sexual orientation, in many cases, they are willing to provide information on the gender of their previous and current partners. Another plausible explanation relates to current relationship status; those currently not in a relationship were more likely to answer "Other" in wave 2016. However, they might have been in a relationship in previous years, thus, can be categorized based on past partnership information.

The cross-table also provides insights on the nature of item nonresponse in self-reports. The vast majority of respondents who provided a non-substantial answer when directly asked about their sexual orientation are classified as heterosexual based on their partnership information. Only one individual reported "Other" but is classified as LGB based on partnership information. Thus, refusing to answer or choosing "Other" seems to be a strategy applied by individuals in opposite-sex partnerships who either feel uncomfortable being asked about their sexuality or do not understand the question and terms (correctly). This is supported by our previous results on mode and interviewer effects on responding "Other".

In a next step, we had a look at potential consequences for applying one or the other measurement strategy when analyzing survey data, for instance, performing comparisons of LGBs and heterosexuals. Table 6 compares the distribution of demographic and socio-economic characteristics across (i) respondents categorized as LGBs by partnership information; (ii) respondents identifying as LGBs by self-reports; (iii) respondents

categorized as heterosexuals by partnership information, and finally (iv) respondents categorized as heterosexuals by self-reports. We only include respondents who report being in a partnership in the wave 2016. Please note that the four groups are not disjunctive, that is, in the case of two consistent measures of sexual orientation, respondents enter both (i) and (ii) or (iii) and (iv); but in the case of inconsistent (= conflicting) measures, respondents enter both (i) and (iv) or (ii) and (iii). Asterisks indicate mean differences between LGBs and heterosexuals at $p < 0.05$, that is, either between columns (i) and (iii) or between columns (ii) and (iv). We investigate whether there are significant mean differences based on one measurement strategy, but not the other.

The first – largely trivial – finding in Table 6 is that respondents whose sexual orientation was coded on the basis of current and previous partnership information are much more likely to currently live in the same household with their partner (85% and 97%). If we draw on self-reported sexual orientation, about 70% of LGBs and 88% of the heterosexuals cohabitate with a partner.

Based on partnership information, LGBs have higher employment rates (77%) than heterosexuals (59%). However, the division of labor in same-sex partnerships differs from opposite-sex partnerships, in that dual-income households are more frequent among same-sex partnerships (Kroh et al. 2017). If we use the self-reported information, we do not find a significant difference in employment by sexual orientation (66% versus 59%) (for effects of the measurement on earning differences by sexual orientation, see Klawitter 2014).

Turning to other characteristics, such as the distribution of gender, highest level of education, and political party identification, both strategies of measuring sexual orientation result in similar estimates. Compared to heterosexuals, not only do LGBs report university entrance degrees ("Abitur", German high school diploma) more frequently, but they also identify more frequently with the Green Party and less frequently with the Christian Democrats.

*Table 6. Descriptives of LGBs and heterosexuals in a partnership in 2016 by measures of sexual orientation.*

| | (i) LGB Partnership N = 136 | (ii) LGB self-report N = 198 | (iii) Heterosexual partnership N = 16,872 | (iv) Heterosexual self-report N = 16,631 |
|---|---|---|---|---|
| Age 17–29 | 0.16 | 0.26 | 0.07* | 0.13* |
| Age 30–45 | 0.45 | 0.38 | 0.33* | 0.33 |
| Age 45–60 | 0.29 | 0.26 | 0.33 | 0.31 |
| Age 60+ | 0.10 | 0.10 | 0.28* | 0.24* |
| Female | 0.60 | 0.61 | 0.61* | 0.52* |
| Partner in HH | 0.85 | 0.69 | 0.97* | 0.88* |
| Employed | 0.77 | 0.66 | 0.59* | 0.59 |
| Municipality < 20K inh. | 0.26 | 0.27 | 0.43* | 0.41* |
| Municipality 20–100K inh. | 0.22 | 0.21 | 0.27 | 0.28* |
| Municipality > 100K inh. | 0.52 | 0.51 | 0.29* | 0.30* |
| Univ. Entrance Degree | 0.50 | 0.45 | 0.37* | 0.38* |
| Support CDU/CSU | 0.17 | 0.16 | 0.39* | 0.38* |
| Support Green Party | 0.36 | 0.34 | 0.14* | 0.15* |

Source. SOEP.v33, Wave 2016. Asterisks indicate mean differences between LGBs and heterosexuals at $p < 0.05$, that is, either between columns (i) and (iii) or between columns (ii) and (iv).

While item nonresponse in self-reported sexual orientation is particularly high among respondents age 65 and older (see Table A2 in the Appendix, Section 6), missing information on the partnership-inferred proxy of sexual orientation decreases by age. This is because younger respondents often lack biographical information on previous partnerships and are still single. Thus, the age distribution of respondents with valid information differs considerably between measures of sexual orientation. In self-reported data, both heterosexuals and LGBs are considerably younger than in the partnership-inferred data. Hence, if differences between heterosexuals and LGBs are present in certain periods of the life cycle, but not in others, the choice of the sexual orientation measure also affects substantive estimates.

## 5. Conclusions

As diversity is an emerging topic in many Western societies and in politics, there is a growing need for data that empirically describes diverse forms of living arrangements, both from academic and public policy perspectives. For instance, EU institutions repeatedly advise member states to monitor the equality of LGBITQ* in various areas of life (see e.g., The European Parliament 2014, "Resolution of 4 February 2014 on the EU roadmap against homophobia and discrimination on grounds of sexual orientation and gender identity"). Exemplifying this trend, official statistics in several countries now report statistics by sexual orientation and gender identity. For example, the U.S. Census publishes data by sexual orientation (U.S. Census Bureau 2019), as does the UK Office for National Statistics (2019), the Australian Bureau of Statistics (2016), and the Federal Statistical Office of Germany (2019), to name a few. However, generalizability of research on sexual orientation is hampered by the lack of consensus on its operationalization. While most researchers presumably would agree that the measurement of self-reported sexual orientation is the preferred option the omission of this measure in many surveys constrains secondary data users to operationalize sexual orientation alternatively using proxy information on the gender of respondents' partner.

The most obvious limitation of this partnership-inferred proxy of sexual orientation is the exclusion of singles and partners living apart, as well as the misclassification of bisexuals. However, the present article shows that implementing self-reports of sexual orientation in surveys also comes with the potential of error. The analyses show that self-reporting is sensitive to mode of data collection and interviewer characteristics. Survey practitioners may choose to rely on interviewers to minimize item nonresponse, but implement the question about sexual orientation within a (computer-assisted) self-interviewing module in order to minimize socially desirable answering behavior (see De Leeuw et al. 2003).

Moreover, we illustrate that the partnership-inferred and the self-reported measures of sexual orientation may result in dissimilar conclusions about the differences between heterosexuals and LGBs. In particular, research linked to with characteristics of partnership, such as occupation, earnings, and social networks is to be interpreted with caution, as is likely to be contingent on the choice of measure of sexual orientation. Therefore, we strongly argue in favour of also collecting data on self-reported sexual orientation in multidisciplinary survey data to enhance the reliability and relevance of LGB research.

Based on our results, there are a number of practical implications when collecting and analyzing survey data on sexual orientation:

- Mode of Data Collection
  - With respect to mode selection, we recommend interviewer-administered self-interviewing by, for instance, CASI mode. Our results suggest that the presence of the interviewer reduces item nonresponse and the privacy of the self-interview increases the prevalence of LGB identification.
  - Question wording and placement need to make sure that respondents understand the question as not only relating to current partnerships.
  - Instead of providing a closed-ended "Other" response option in surveying sexual orientation, which seems to be used as a substitute for refusals by some respondents, we recommend using an open-ended response option with follow-up write-ins.
- Interviewer Effects
  - The large interviewer effects reported in the article point to the necessity of extensive interviewer training. Interviewers should be specifically informed about the importance of the question and how to handle potential concerns, uncertainties and discomfort of respondents (and themselves). Also, interviewer monitoring during fieldwork may be used to identify interviewers who struggle with asking the question appropriately.
  - We suggest acknowledging interviewer effects in the analysis of the survey data by estimating multi-level regression models. This prevents researchers from underestimating standard errors and minimizes the likelihood of type 1 errors, that is, the rejection of a true null hypothesis.
- Self-reported versus partnership-inferred sexual orientation
  - Wherever possible, researchers should obtain self-reports on sexual orientation rather than information on partnership constellation only. Self-reports allow to address the general population instead of the subpopulation of cohabitating couples, and they allow to classify bisexual (and e.g., pansexual) respondents. Also, from the perspective of research ethics, directly probing sensitive information, including an open-ended response option appears superior to inferring the information from other sources.
  - In case both self-reports and partnership information is available, we suggest combining both sources of information, for instance, by analyzing current and previous partnerships for respondents that refused to provide self-reports. In the case of the SOEP sample, combining both ways of measuring sexual orientation allows classifying 97% of all adult respondents as either LGB or heterosexual. Although partnership-inferred information used as a replacement for self-reported information may be marred by misclassification error, the combination of data lends itself for estimating the bounds or error. Alternatively, we may use partnership-inferred information and self-reported information with a large set of additional variables within a multiple-imputation framework, replacing missing information. Again, the partnership-inferred information will be a strong predictor of sexual orientation and this approach lends itself for assessing the uncertainty of combining information.

## 6. Appendix

*Table A1. Composition of respondents across modes.*

| | Face-to-face | | Self-administered | |
|---|---|---|---|---|
| Variable | N | % | n | % |
| Third person present during interview | | | | |
| No information | 9,382 | 53 | 5,266 | 78 |
| Partner | 4,314 | 25 | 1,151 | 17 |
| Other | 3,587 | 20 | 280 | 4 |
| No third person | 262 | 2 | 45 | 1 |
| Respondent Panel Experience | | | | |
| 1–2 years | 1,634 | 9 | 286 | 4 |
| 3–5 years | 5,022 | 29 | 501 | 7 |
| 6–10 years | 6,285 | 36 | 1,528 | 23 |
| 11+ years | 4,604 | 26 | 4,427 | 66 |
| Subsample | | | | |
| Early cross-sectional | 8,452 | 48 | 3,479 | 52 |
| Migrants | 4,085 | 23 | 304 | 5 |
| Families | 3,869 | 22 | 1,360 | 20 |
| Other | 1,139 | 7 | 1,599 | 23 |
| Respondent Sex | | | | |
| Male | 8,051 | 46 | 3,068 | 46 |
| Female | 9,494 | 54 | 3,674 | 54 |
| Respondent Age | | | | |
| <25 | 1,888 | 11 | 775 | 12 |
| 26–35 | 2,602 | 15 | 766 | 11 |
| 36–50 | 5,689 | 32 | 1,968 | 29 |
| 51–65 | 3,856 | 22 | 2,002 | 30 |
| 66+ | 3,510 | 20 | 1,231 | 18 |
| Household size | | | | |
| Single | 1,617 | 9 | 381 | 6 |
| 2 | 4,706 | 27 | 1,410 | 21 |
| 3 | 3,241 | 18 | 1,316 | 20 |
| 4 | 4,144 | 23 | 1,866 | 28 |
| 5 | 2,195 | 13 | 1,124 | 17 |
| 6+ | 1,672 | 10 | 645 | 10 |
| Region | | | | |
| West Germany | 14,242 | 81 | 5,010 | 74 |
| East Germany | 3,303 | 19 | 1,732 | 26 |

*Table A2. Full model – explaining interviewer effects.*

| | Nonresponse (y/n) | | Rep. "other" (y/n) | |
|---|---|---|---|---|
| Variable (interviewer) | OR | 95% CI | OR | 95% CI |
| Gender | | | | |
| Male (ref.) | – | – | – | – |
| Female | 0.80 | 0.47–1.38 | 0.66 | 0.32–1.33 |
| Age | 1.01 | 0.98–1.04 | 1.09*** | 1.05–1.15 |
| Education | | | | |
| Basic (ref.) | – | – | – | – |
| Secondary | 0.70 | 0.34–1.42 | 0.32* | 0.13–0.77 |
| Tertiary | 0.73 | 0.34–1.55 | 0.31* | 0.12–0.81 |
| Work experience (years) | | | | |
| 1–10 (ref.) | – | – | – | – |
| 11–20 | 1.50 | 0.82–2.76 | 0.75 | 0.35–1.62 |
| 21+ | 2.82** | 1.38–5.74 | 0.57 | 0.21–1.58 |
| Number of personal interviews in wave 2016 | | | | |
| 1–49 | – | – | – | – |
| 50–99 | 0.43** | 0.24–0.77 | 0.69 | 0.33–1.46 |
| 100+ | 0.37* | 0.17–0.79 | 0.94 | 0.36–2.40 |
| Personality-score (Big five)[1] | | | | |
| Openness | 1.04 | 0.79–1.35 | 1.12 | 0.78–1.64 |
| Conscientiousness | 1.37 | 0.96–1.95 | 1.41 | 0.89–2.21 |
| Extraversion | 0.91 | 0.65–1.26 | 1.61* | 1.03–2.50 |
| Agreeableness | 0.98 | 0.70–1.39 | 1.42 | 0.91–2.21 |
| Neuroticism | 1.19 | 0.96–1.49 | 1.11 | 0.84–1.48 |
| Party identification | | | | |
| No party (ref.) | – | – | – | – |
| SPD | 0.72 | 0.33–1.57 | 0.62 | 0.22–1.75 |
| CDU/CSU | 0.59 | 0.31–1.14 | 1.47 | 0.65–3.34 |
| Greens | 0.65 | 0.23–1.80 | 0.37 | 0.08–1.67 |
| The Left | 0.44 | 0.15–1.35 | 1.71 | 0.44–6.71 |
| Other | 0.76 | 0.24–2.52 | 0.98 | 0.21–4.48 |
| Third person present during interview | | | | |
| No information (ref.) | – | – | – | – |
| Partner | 0.71 | 0.44–1.15 | 0.63 | 0.38–1.04 |
| Other | 0.61 | 0.37–1.02 | 0.82 | 0.47–1.45 |
| No third person | 0.52** | 0.34–0.80 | 0.38*** | 0.24–0.61 |

*Table A2. Continued.*

| Variable (interviewer) | Nonresponse (y/n) | | Rep. "other" (y/n) | |
|---|---|---|---|---|
| | OR | 95% CI | OR | 95% CI |
| Time of interview within individual interviewer workload | | | | |
| First half (ref.) | – | – | – | – |
| Second half | 1.11 | 0.88–1.39 | 0.56*** | 0.43–0.73 |
| Subsample | | | | |
| Cross-sectional (ref.) | – | – | – | – |
| Migrants | 1.99*** | 1.36–2.93 | 1.44 | 0.88–2.34 |
| Families | 0.64* | 0.42–0.96 | 3.79*** | 2.54–5.65 |
| Other | 0.82 | 0.49–1.37 | 1.14 | 0.66–1.97 |
| German region | | | | |
| North | – | – | – | – |
| East | 1.74 | 0.74–4.10 | 0.60 | 0.21–1.68 |
| South | 1.97 | 0.85–4.59 | 0.92 | 0.34–2.54 |
| West | 1.98 | 0.90–4.34 | 0.80 | 0.32–2.01 |
| Municipality size | | | | |
| <2,000 (ref.) | – | – | – | – |
| 2,000–5,000 | 0.65 | 0.39–1.09 | 0.88 | 0.46–1.70 |
| 5,000–20,000 | 0.76 | 0.49–1.16 | 0.97 | 0.56–1.79 |
| 20,000–50,000 | 0.59* | 0.36–0.95 | 0.81 | 0.44–1.51 |
| 50,000–100,000 | 0.61 | 0.34–1.08 | 0.63 | 0.31–1.28 |
| 100,000–500,000 | 0.94 | 0.57–1.54 | 0.70 | 0.37–1.34 |
| 500,000+ | 0.58 | 0.33–1.02 | 0.56 | 0.26–1.20 |
| Respondent panel experience | | | | |
| 1–2 years (ref.) | – | – | – | – |
| 3–5 years | 0.77 | 0.52–1.14 | 0.65 | 0.41–1.03 |
| 6–10 years | 0.81 | 0.50–1.31 | 1.11 | 0.65–1.88 |
| 11+ years | 1.05 | 0.63–1.73 | 0.91 | 0.50–1.64 |
| Respondent sex | | | | |
| male (ref.) | – | – | – | – |
| female | 1.30** | 1.07–1.59 | 1.21 | 0.97–1.52 |
| Respondent age | | | | |
| <25 (ref.) | – | – | – | – |
| 26–35 | 0.81 | 0.50–1. 31 | 0.57* | 0.33–0.97 |
| 36–50 | 0.96 | 0.62–1.47 | 0.98 | 0.62–1.54 |
| 51–65 | 1.70* | 1.09–2.64 | 1.36 | 0.84–2.22 |
| 66+ | 2.58*** | 1.60–4.16 | 3.29*** | 1.92–5.65 |

*Table A2. Continued.*

| Variable (interviewer) | Nonresponse (y/n) | | Rep. "other" (y/n) | |
|---|---|---|---|---|
| | OR | 95% CI | OR | 95% CI |
| Respondent education | | | | |
| basic (ref.) | – | – | – | – |
| secondary | 1.06 | 0.84–1.35 | 1.01 | 0.77–1.32 |
| tertiary | 0.60** | 0.43–0.82 | 0.74 | 0.53–1.04 |
| In a relationship? | | | | |
| No (ref.) | – | – | – | – |
| Yes | 0.29*** | 0.22–0.37 | 0.40*** | 0.30–0.55 |
| Household size | | | | |
| Single (ref.) | – | – | – | – |
| 2 | 0.91 | 0.64–1.31 | 0.81 | 0.52–1.26 |
| 3 | 0.94 | 0.62–1.40 | 0.92 | 0.57–1.46 |
| 4 | 1.02 | 0.67–1.55 | 0.84 | 0.51–1.39 |
| 5 | 1.31 | 0.83–2.09 | 1.10 | 0.65–1.89 |
| 6+ | 1.14 | 0.70–1.86 | 0.87 | 0.51–1.50 |
| Constant | 0.00** | 0.00–0.24 | 0.00*** | 0.00–0.00 |
| $n_{respondents}$ | 14,521 | | 13,389 | |
| $n_{interviewers}$ | 432 | | 401 | |

*p < .05, **p < .01, ***p < .001.
[1]Range: 1 to 5 with lower values representing less open, agreeable, and so on.
Multilevel logistic regression with respondents nested in interviewers.

## 7. References

Aicken, C.R.H., C.H. Mercer, and J.A. Cassell. 2013. "Who reports absence of sexual attraction in Britain? Evidence from national probability surveys." *Psychology and Sexuality* 4(2): 121–135. DOI: https://doi.org/10.1080/19419899.2013.774161.

Antecol, H., A. Jong, and M. Steinberger. 2008. "The Sexual Orientation Wage Gap: The Role of Occupational Sorting and Human Capital." *Industrial and Labor Relations Review* 61(4): 518–543. DOI: https://doi.org/10.1177/001979390806100405.

Australian Bureau of Statistics. 2016. Same-Sex Couples in Australia, 2016. Available at: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/2071.0Main+Features852016?OpenDocument (accessed October 2019).

Bates, N. and T.J. DeMaio. 2013. "Measuring same-sex relationships." *Contexts* 12: 66–69. DOI: https://doi.org/10.1177/1536504213476251.

Black, D.A. 2000. "Demographics of the Gay and Lesbian Population in the United States: Evidence from Available Systematic Data Sources." *Demography* 37(2): 139–154. DOI: https://doi.org/10.2307/2648117.

Black, D.A., S.G. Sanders, and L.J. Taylor. 2007. "The Economics of Lesbian and Gay Families." *The Journal of Economic Perspectives* 21(2): 53–70. DOI: https://doi.org/10.1257/jep.21.2.53.

Booker, C.L., G. Riger, and J.B. Unger. 2017. "Sexual orientation health inequality: Evidence from Understanding Society, the UK Longitudinal Household Study." *Preventive Medicine* 101: 126–132. DOI: https://doi.org/10.1016/j.ypmed.2017.06.010.

Campanelli, P. and C. O'Muricheartaigh. 1999. "Interviewers, Interviewer Continuity, and Panel Survey Nonresponse." *Quality & Quantity* 33(1): 59–76. DOI: https://doi.org/10.1023/a:1004357711258.

Chandra, A., C.E. Copen, and W.D. Mosher. 2013. "Sexual Behavior, Sexual Attraction, and Sexual Identity in the United States: Data from the 2006–2010 National Survey of Family Growth." In *International Handbook on the Demography of Sexuality*, vol 5, edited by A. Baumle, 45–66. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-94-007-5512-3_4.

Cortina, C., and P. Festy. 2014. *Identification of same-sex couples and families in censuses, registers and surveys*. Families and Societies Working Paper Series 8. Available at: http://www.familiesandsocieties.eu/wp-content/uploads/2014/12/WP8CortinaFesty2014.pdf (accessed October 2019).

Dahlhamer, J.M., A.M. Galinksy, S.S. Joestl, and B.W. Ward. 2014. "Sexual orientation in the National Health Interview Survey: A quality assessment." *Vital Health Stat* 2(169) Available at: https://www.researchgate.net/publication/269713986_Sexual_Orientation_in_the_2013_National_Health_Interview_Survey_A_Quality_Assessment (accessed November 2019).

De Leeuw, E., J. Hox, and S. Kef. 2003. "Computer-Assisted Self-Interviewing Tailored for Special Populations and Topics." *Field Methods* 15(3): 223–251. DOI: https://doi.org/10.1177/1525822x03254714.

DeMaio, T.J., N. Bates, and M. O'Connell. 2013. "Exploring Measurement Error Issues in Reporting of Same-Sex Couples." *Public Opinion Quarterly* 77(1): 145–158. DOI: https://doi.org/10.1093/poq/nfs066.

Dillman, D.A., G. Phelps, R. Tortora, K. Swift, J. Kohrell, J. Berck, and B.L. Messer. 2009. "Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet." *Social Science Research* 38: 1–18. DOI: https://doi.org/10.1016/j.ssresearch.2008.03.007.

Durrant, G.B. and J. D'Arrigo. 2014. "Doorstep Interactions and Interviewer Effects on the Process Leading to Cooperation or Refusal." *Sociological Methods & Research* 43(3): 490–518. DOI: https://doi.org/10.1177/0049124114521148.

Federal Statistical Office of Germany. 2019. Gleichgeschlechtlichen Paare (darunter: gleichgeschlechtliche Ehepaare und eingetragene Lebenspartnerschaften). Available at: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Haushalte-Familien/Tabellen/3-4-gleichgeschlechtliche-lebensgemeinschaften.html?nn=209096 (accessed October 2019).

Festy, P. 2007. "Enumerating same-sex couples in censuses and population registers." *Demographic Research* 17: 339–368. DOI: https://doi.org/10.4054/demres.2007.17.12.

Frederiksen-Goldsen, K.I. and H.-J. Kim. 2014. "Count Me In: Response to Sexual Orientation Measures Among Older Adults." *Research on Aging* 37(5): 464–480. DOI: https://doi.org/10.1177/0164027514542109.

Galupo, M.P., E. Lomash, and R.C. Mitchell. 2016. "All of My Lovers Fit Into This Scale: Sexual Minority Individuals' Responses to Two Novel Measures of Sexual Orientation." *Journal of Homosexuality* 64: 145–165. DOI: https://doi.org/10.1080/00918369.2016.1174027.

Gates, G.J. 2011. *How many people are lesbian, gay, bisexual and transgender?* Los Angeles: The Williams Institute. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/Gates-How-Many-People-LGBT-Apr-2011.pdf (accessed October 2019).

Geary, R.S., C. Tanton, B. Erens, S. Clifton, P. Prah, K. Wellings, K.R. Mitchell, J. Datta, K. Gravningen, E. Fuller, A.M. Johnson, P. Sonnenberg, and C.H. Mercer. 2018. "Sexual identity, attraction and behaviour in Britain: The implications of using different dimensions of sexual orientation to estimate the size of sexual minority populations and inform public health interventions." *PLoS ONE* 13(1). DOI: https://doi.org/10.1371/journal.pone.0189607.

Goebel, J., M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp. 2018. "The German Socio-Economic Panel (SOEP)." *Journal of Economic and Statistics* 239(2): 345–360. DOI: https://doi.org/10.1515/jbnst-2018-0022.

Groves, R.M. 2004. *Survey Error and Survey Costs*. Hoboken, NJ: John Wiley & Sons.

Hank, K. and M. Wetzel. 2018. "Same-sex relationship experiences and expectations regarding partnership and parenthood." *Demographic Research* 39: 701–718. DOI: https://doi.org/10.4054/demres.2018.39.25.

Haseldon, L. and T. Joloza. 2009. "Measuring Sexual Identity: A Guide for Researchers." Office for National Statistics, Newport UK. Available at: https://osservatoriogender.it/wp-content/uploads/2016/10/sexualidentityuserguidefinal_tcm77-181188.pdf (accessed October 2019).

Heckman, J. 1976. "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." *Annals of Economic and Social Measurement* 5: 475–492. Available at: https://ideas.repec.org/h/nbr/nberch/10491.html (accessed November 2019).

Heckman, J. 1979. "Sample selection bias as a specification error." *Econometrica* 47(3): 153–161. DOI: https://doi.org/10.2307/1912352.

Hilgert, L. M. Kroh, and D. Richter 2016. "The effect of face-to-face interviewing on personality measurement." *Journal of Research in Personality* 63: 133–136. DOI: https://doi.org/10.1016/j.jrp.2016.05.006.

Jans, M., J. Viana, D. Grant, S.D. Cochran, A.C. Lee, and N.A. Ponce. 2015. "Trends in sexual orientation missing data over a decade of the California Health Interview Survey." *American Journal of Public Health* 105(5): e43–e50. DOI: https://doi.org/10.2105/ajph.2014.302514.

Kish, Leslie. 1962. "Studies of Interviewer Variance for Attitudinal Variables." *Journal of the American Statistical Association* 57(297): 92–115. DOI: https://doi.org/10.1080/01621459.1962.10482153.

Klawitter, M. 2011. "Multilevel Analysis of the Effects of Antidiscrimination Policies on Earnings by Sexual Orientation." *Journal of Policy Analysis and Management* 30(2): 334–358. DOI: https://doi.org/10.1002/pam.20563.

Klawitter, M. 2014. "Meta-Analysis of the Effects of Sexual Orientation on Earnings." *Industrial Relations: A Journal of Economy and Society* 54(1): 4–32. DOI: https://doi.org/10.1111/irel.12075.

Kreider, R.M. and D.A. Lofquist. 2015. "Matching Survey Data with Administrative Records to Evaluate Reports of Same-Sex Married Couple Households." *SEHSD Working Paper*, 2014-36. U.S. Census Bureau: Washington, DC.

Kroh, M., S. Kühne, C. Kipp, and D. Richter. 2017. "Income, social support networks, life satisfaction: lesbians, gays, and bisexuals in Germany." *DIW Economic Bulletin* 2017: 335–345. Available at: https://diw.de/documents/publikationen/73/diw_01.c.563931.de/diw_econ_bull_2017-33-2.pdf (accessed October 2019).

Kroh, M., S. Kühne, R. Siegers, and V. Belcheva. 2018. "SOEP-Core: Documentation of Sample Sizes and Panel Attrition (1984 until 2016)." *SOEP Survey Papers* 480. Available at: https://www.diw.de/documents/publikationen/73/diw_01.c.579464.de/diw_ssp0480.pdf (accessed October 2019).

Krosnick, J.A., S. Narayan, and W.R. Smith. 1996. "Satisficing in surveys: Initial evidence." *New Directions for Evaluation* 1996(70): 29–44. DOI: https://doi.org/10.1002/ev.1033.

Kuha, J., C. Skinner, and S. Butt. 2014. *The effect of interviewer probing on item nonresponse and measurement error in cross-national surveys*. Presentation held at the 2014 American Association for Public Opinion Research Conference, May 15–18, 2014, Anaheim, California, U.S.A. Available at: https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/Abstract_Book_FNL.pdf (accessed November 2019).

Kühne, S. 2018. "From Strangers to Acquaintances? Interviewer Continuity and Socially Desirable Responses in Panel Surveys." *Survey Research Methods* 12(2): 121–146. DOI: https://doi.org/10.18148/srm/2018.v12i2.7299.

Leppel, K. 2009. "Labour Force Status and Sexual Orientation." *Economica* 76(301): 197–207. DOI: https://doi.org/10.1111/j.1468-0335.2007.00676.x.

Lewis, J.M., N. Bates, and M. Streeter. 2019. *Measuring Same-Sex Couples: The What And Who of Misreporting on Relationship And Sex*. U.S. Census Bureau. Available at: https://census.gov/content/dam/Census/library/working-papers/2015/demo/SEHSD-WP2015-12.pdf (accessed October 2019).

Lengerer, A. 2017. *Quality of Official Data on Cohabiting Same-Sex Couples in Germany*. Presentation held at the 2017 European Survey Research Association Conference, July 17th–21st 2017, Lisbon, Portugal.

Liu, H., C. Reczek, and D. Brown. 2013. "Same-sex Cohabitors and Health: the role of race-ethnicity, gender, and socioeconomic status." *Journal of Health and Social Behavior* 54(1): 25–45. DOI: https://doi.org/10.1177/0022146512468280.

Lofquist, D., T. Lugaila, M. O'Connell, and S. Feliz. 2012. "Households and Families: 2010." *Census Briefs*. C2010BR-14. Available at: https://www.census.gov/prod/cen2010/briefs/c2010br-14.pdf (accessed November 2019).

McCrae, R.R. and O.P. John. 1992. "An Introduction to the Five-Factor Model and Its Applications." *Journal of Personality* 60(2): 175–215. DOI: https://doi.org/10.1111/j.1467-6494.1992.tb00970.x.

Meyer, I.H. and P.A. Wilson. 2009. "Sampling lesbian, gay, and bisexual populations." *Journal of Counseling Psychology* 56(1): 23–31. DOI: https://doi.org/10.1037/a0014587.

Moleiro, C. and N. Pinto. 2015. "Sexual orientation and gender identity: review of concepts, controversies and their relation to psychopathology classification systems." *Frontiers in psychology* 6: 1511. DOI: https://doi.org/10.3389/fpsyg.2015.01511.

Puhani, P. 2000. "The Heckman correction for sample selection and its critique." *Journal of economic surveys* 14(1): 53–68. DOI: https://doi.org/10.1111/1467-6419.00104.

Rasbash, J. and H. Goldstein. 1994. "Efficient Analysis of Mixed Hierarchical and Cross-Classified Random Structures Using a Multilevel Model." *Journal of Educational and Behavioral Statistics* 19(4): 337–350. DOI: https://doi.org/10.3102/10769986019004337.

Richter, D. and J. Schupp. 2015. "The SOEP Innovation Sample (SOEP IS)." *Journal of Contextual Economics – Schmollers Jahrbuch* 135(3): 389–399. DOI: https://doi.org/10.3790/schm.135.3.389.

Schnell, R. and F. Kreuter. 2005. "Separating Interviewer and Sampling-Point Effects." *Journal of Official Statistics* 21(3): 389–41020 Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/separating-interviewer-and-sampling-point-effects.pdf (accessed November 2019).

Schwarz, N., F. Strack, H.-J. Hippler, and G. Bishop. 1991. "The impact of administration mode on response effects in survey measurement." *Applied Cognitive Psychology* 5(3): 193–212. DOI: https://doi.org/10.1002/acp.2350050304.

Sell, R.L. 2007. "Defining and Measuring Sexual Orientation for Research." In: *The Health of Sexual Minorities*, edited by I.H. Meyer and M.E. Northridge. Boston, MA: Springer. DOI: https://doi.org/10.1007/978-0-387-31334-4_14.

The European Parliament. 2014 "Resolution of 4 February 2014 on the EU roadmap against homophobia and discrimination on grounds of sexual orientation and gender identity". Available at: http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0062+0+DOC+XML+V0//EN (accessed October 2019)

Tourangeau, R. and T. Yan. 2007. "Sensitive Questions in Surveys." *Psychological Bulletin* 133(5): 859–883. DOI: https://doi.org/10.1037/0033-2909.

Uhrig, N. 2014. *An Examination of Poverty and Sexual Orientation in the UK*. Essex: Institute for Social and Economic Research. Available at: https://iser.essex.ac.uk/research/publications/working-papers/iser/2014-02.pdf (accessed October 2019).

UK Office for National Statistics. 2019. Sexual Identity lesbian, gay and bisexual population by Occupation by Country. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/sexuality/datasets/sexualidentitylesbiangayandbisexualpopulationbyoccupationbycountry (accessed October 2019).

U.S. Census Bureau. 2019. Same-Sex Couples. Available at: https://www.census.gov/topics/families/same-sex-couples/data/tables.html (accessed October 2019).

West, B.T. and K. Olson. 2010. "How Much Interviewer Variance Is Really Nonresponse Error Variance?" *Public Opinion Quarterly* 74(5): 1004–1026. DOI: https://doi.org/10.1093/poq/nfq061.

Williams Institute. 2009. *Best Practices for Asking Questions about Sexual Orientation on Surveys?* Los Angeles: The Williams Institute. Available at: https://www.williamsinstitute.law.ucla.edu/wp-content/uploads/SMART-FINAL-Nov-2009.pdf (accessed October 2019).

Wolff, M., B. Wells, C. Ventura-DiPersia, A. Renson, and C. Grov. 2016. "Measuring Sexual Orientation: A Review and Critique of U.S Data Collection Efforts and Implications for Health Policy." *The Journal of Sex Research* 54: 507–531. DOI: https://doi.org/10.1080/00224499.2016.1255872.

Wooden, M. 2014. *The Measurement of Sexual Identity in Wave 12 of the HILDA Survey (and associations with mental health and earnings)*. Melbourne: Melbourne Institute for Applied Economic and Social Research. Available at: https://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/hilda-discussion-papers/hdps114.pdf (accessed October 2019).

Received November 2019
Revised June 2019
Accepted October 2019